OMI Research Newsletter – December 2023

The Oxford-Man Institute of Quantitative Finance (OMI) is a world-leading centre for interdisciplinary research in financial markets. The OMI addresses fundamental problems in quantitative finance with a strong focus on machine learning and data driven models.

*The **OMI Research Newsletter** is a curated summary that covers the most recent research, insights, and tools related to financial decision making.*

Highlights

News. The OMI is launching the OMI Finance Seminar, starting January 2024. The aim of the OMI Finance Seminar is to discuss and promote research, theoretical and applied, that addresses the emerging challenges in finance. The seminar will focus on recent and cutting-edge contributions to topics that can help lay the foundations for the future financial landscape including economics, investment, market microstructure, monetary policy, decentralised finance, and financial technology. The seminar will be on Zoom. To join us and for more information please follow this link here to send a request.

The Machine Learning and Financial Econometrics Conference will take place on 12 and 13 December 2023 at the Oxford-Man Institute. Click here for a schedule and to register.

Research focus. Noteworthy contributions in this month's edition include works from the OMI and Imperial College London. Cartea, Chang and García-Arenas (2023) propose a dynamic model of the limit order book and derive conditions to test if a trading algorithm will learn to manipulate the order book. Achitouv, Gorduza and Jacquier (2023) provide an overview of Natural Language Processing techniques for financial regulation. Stok and Bilokon (2023) use a hybrid architecture with neural networks and particle filtering to calculate volatility.

Machine Learning & Statistics

From Deep Filtering to Deep Econometrics

R. STOK, P. BILOKON
IMPERIAL COLLEGE LONDON

Calculating true volatility is an essential task for option pricing and risk management. However, it is made difficult by market microstructure noise. Particle filtering has been proposed to solve this problem as it favorable statistical properties, but relies on assumptions about underlying market dynamics. Machine learning methods have also been proposed but lack interpretability, and often lag in performance. In this paper we implement the SV-PF-RNN: a hybrid neural network and particle filter architecture. Our SV-PF-RNN is designed specifically with stochastic volatility estimation in mind. We then show that it can improve on the performance of a basic particle filter.

Quantum-inspired nonlinear Galerkin ansatz for high-dimensional HJB equations

C. SUN, A. COHEN, J. STOKES, S. VEERAPANENI
UNIVERSITY OF MICHIGAN

Neural networks are increasingly recognized as a powerful numerical solution technique for partial differential equations (PDEs) arising in diverse scientific computing domains, including quantum many-body physics. In the context of time-dependent PDEs, the dominant paradigm involves casting the approximate solution in terms of stochastic minimization of an objective function given by the norm of the PDE residual, viewed as a function of the neural network parameters. Recently, advancements have been made in the direction of an alternative approach which shares aspects of nonlinearly parametrized Galerkin methods and variational quantum

Monte Carlo, especially for high-dimensional, time-dependent PDEs that extend beyond the usual scope of quantum physics. This paper is inspired by the potential of solving Hamilton-Jacobi-Bellman (HJB) PDEs using Neural Galerkin methods and commences the exploration of nonlinearly parametrized trial functions for which the evolution equations are analytically tractable. As a precursor to the Neural Galerkin scheme, we present trial functions with evolution equations that admit closed-form solutions, focusing on time-dependent HJB equations relevant to finance.

Natural Language Processing for Financial Regulation

I. Achitouv, D. Gorduza, A. Jacquier
Imperial College London, Oxford Man Institute

This article provides an understanding of Natural Language Processing techniques in the framework of financial regulation, more specifically in order to perform semantic matching search between rules and policy when no dataset is available for supervised learning. We outline how to outperform simple pre-trained sentences-transformer models using freely available resources and explain the mathematical concepts behind the key building blocks of Natural Language Processing.

Enhancing Financial Sentiment Analysis via Retrieval Augmented Large Language Models

B. Zhang, H. Yang, T. Zhou, A. Babar, X-Y. Liu
The University of Adelaide, Columbia University, Brown University

Financial sentiment analysis is critical for valuation and investment decision-making. Traditional NLP models, however, are limited by their parameter size and the scope of their training datasets, which hampers their generalization capabilities and effectiveness in this field. Recently, Large Language Models (LLMs) pre-trained on extensive corpora have demonstrated superior performance across various NLP tasks due to their commendable zero-shot abilities. Yet, directly applying LLMs to financial sentiment analysis presents challenges: The discrepancy between the pre-training objective of LLMs and predicting the sentiment label can compromise their predictive performance. Furthermore, the succinct nature of financial news, often devoid of sufficient context, can significantly diminish the reliability of LLMs' sentiment analysis. To address these challenges, we introduce a retrieval-augmented LLMs framework for financial sentiment analysis. This framework includes an instruction-tuned LLMs module, which ensures LLMs behave as predictors of sentiment labels, and a retrieval-augmentation module which retrieves additional context from reliable external sources. Benchmarked against traditional models and LLMs like ChatGPT and LLaMA, our approach achieves 15% to 48% performance gain in accuracy and F1 score.

Conformal Prediction for Time Series with Modern Hopfield Networks

A. Auer, M. Gauch, D. Klotz, S. Hochreiter
Johannes Kepler University, Institute for Machine Learning

To quantify uncertainty, conformal prediction methods are gaining continuously more interest and have already been successfully applied to various domains. However, they are difficult to apply to time series as the autocorrelative structure of time series violates basic assumptions required by conformal prediction. We propose HopCPT, a novel conformal prediction approach for time series that not only copes with temporal structures but leverages them. We show that our approach is theoretically well justified for time series where temporal dependencies are present. In experiments, we demonstrate that our new approach outperforms state-of-the-art conformal prediction methods on multiple real-world time series datasets from four different domains.

Implicit Maximum a Posteriori Filtering via Adaptive Optimization

G. M. Bencomo, J. C. Snell, T. L. Griffiths
Princeton University

Bayesian filtering approximates the true underlying behavior of a time-varying system by inverting an explicit generative model to convert noisy measurements into state estimates. This process typically requires either storage, inversion, and multiplication of large matrices or Monte Carlo estimation, neither of which are practical in high-dimensional state spaces such as the weight spaces of artificial neural networks. Here, we frame the standard Bayesian filtering problem as optimization over a time-varying objective. Instead of maintaining matrices for the filtering equations or simulating particles, we specify an optimizer that defines the Bayesian filter implicitly. In the linear-Gaussian setting, we show that every Kalman filter has an equivalent formulation using K steps of gradient descent. In the nonlinear setting, our experiments demonstrate that our framework results in filters that are effective, robust, and scalable to high-dimensional systems, comparing well against the standard toolbox of Bayesian filtering solutions. We suggest that it is easier to fine-tune an optimizer than it is to specify the correct filtering equations, making our framework an attractive option for high-dimensional filtering problems.

Signatures Meet Dynamic Programming: Generalizing Bellman Equations for Trajectory Following

ANONYMOUS AUTHORS
ANONYMOUS

Path signatures have been proposed as a powerful representation of paths that efficiently captures the path's analytic and geometric characteristics, having useful algebraic properties including fast concatenation of paths through tensor products. Signatures have recently been widely adopted in machine learning problems for time series analysis. In this work we establish connections between value functions typically used in optimal control and intriguing properties of path signatures. These connections motivate our novel control framework with signature transforms that efficiently generalizes the Bellman equation to the space of trajectories. We analyze the properties and advantages of the framework, termed signature control. In particular, we demonstrate that (i) it can naturally deal with varying/adaptive time steps; (ii) it propagates higher-level information more efficiently than value function updates; (iii) it is robust to dynamical system misspecification over long rollouts. As a specific case of our framework, we devise a model predictive control method for path tracking. This method generalizes integral control, being suitable for problems with unknown disturbances. The proposed algorithms are tested in simulation, with differentiable physics models including typical control and robotics tasks such as point-mass, curve following for an ant model, and a robotic manipulator.

Deep self-consistent learning of local volatility

Z. WANG, N. PRIVAULT, C. GUET
NANYANG TECHNOLOGICAL UNIVERSITY

We present an algorithm for the calibration of local volatility from market option prices through deep self-consistent learning, by approximating both market option prices and local volatility using deep neural networks, respectively. Our method uses the initial-boundary value problem of the underlying Dupire's partial differential equation solved by the parameterized option prices to bring corrections to the parameterization in a self-consistent way. By exploiting the differentiability of the neural networks, we can evaluate Dupire's equation locally at each strike-maturity pair; while by exploiting their continuity, we sample strike-maturity pairs uniformly from a given domain, going beyond the discrete points where the options are quoted. Moreover, the absence of arbitrage opportunities are imposed by penalizing an associated loss function as a soft constraint. For comparison with existing approaches, the proposed method is tested on both synthetic and market option prices, which shows an improved performance in terms of reduced interpolation and reprice errors, as well as the smoothness of the calibrated local volatility. An ablation study has been performed, asserting the robustness and significance of the proposed method.

Maximizing Portfolio Predictability with Machine Learning

M. PINELIS, D. RUPPERT
CORNELL UNIVERSITY

We construct the maximally predictable portfolio (MPP) of stocks using machine learning. Solving for the optimal constrained weights in the multi-asset MPP gives portfolios with a high monthly coefficient of determination, given the sample covariance matrix of predicted return errors from a machine learning model. Various models for the covariance matrix are tested. The MPPs of S&P 500 index constituents with estimated returns from Elastic Net, Random Forest, and Support Vector Regression models can outperform or underperform the index depending on the time period. Portfolios that take advantage of the high predictability of the MPP's returns and employ a Kelly criterion style strategy consistently outperform the benchmark.

Robust and Conjugate Gaussian Process Regression

M. ALTAMIRANO, F-X. BRIOL, J. KNOBLAUCH
UNIVERSITY COLLEGE LONDON

To enable closed form conditioning, a common assumption in Gaussian process (GP) regression is independent and identically distributed Gaussian observation noise. This strong and simplistic assumption is often violated in practice, which leads to unreliable inferences and uncertainty quantification. Unfortunately, existing methods for robustifying GPs break closed-form conditioning, which makes them less attractive to practitioners and significantly more computationally expensive. In this paper, we demonstrate how to perform provably robust and conjugate Gaussian process (RCGP) regression at virtually no additional cost using generalised Bayesian inference. RCGP is particularly versatile as it enables exact conjugate closed form updates in all settings where standard GPs admit them. To demonstrate its strong empirical performance, we deploy RCGP for problems ranging from Bayesian optimisation to sparse variational Gaussian processes.

Finance and Economics

Spoofing and Manipulating Order Books with Learning Algorithms

Á. Cartea, P. Chang, G. García-Arenas

Oxford-Man Institute

We propose a dynamic model of the limit order book to derive conditions to test if a trading algorithm will learn to manipulate the order book. Our results show that as a market maker becomes more tolerant to bearing inventory risk, the learning algorithm will find optimal strategies that manipulate the book more frequently. Manipulation occurs to induce mean reversion in inventory to an optimal level and to execute round-trip trades with limit orders at a higher probability than was otherwise likely to occur; spoofing is a special case when the market maker prefers that manipulative limit orders are not filled. The conditions are tested with order book data from Nasdaq and we show that market conditions are conducive for an algorithm to learn to manipulate the order book. Finally, when two market makers use learning algorithms to trade, their algorithms can learn to coordinate their manipulation.

Algorithmic Persuasion Through Simulation: Information Design in the Age of Generative AI

K. Harris, N. Immorlica, B. Lucier, A. Slivkins

Carnegie Mellon University

How can an informed sender persuade a receiver, having only limited information about the receiver's beliefs? Motivated by research showing generative AI can simulate economic agents, we initiate the study of information design with an oracle. We assume the sender can learn more about the receiver by querying this oracle, e.g., by simulating the receiver's behavior. Aside from AI motivations such as general-purpose Large Language Models (LLMs) and problem-specific machine learning models, alternate motivations include customer surveys and querying a small pool of live users. Specifically, we study Bayesian Persuasion where the sender has a second-order prior over the receiver's beliefs. After a fixed number of queries to an oracle to refine this prior, the sender commits to an information structure. Upon receiving the message, the receiver takes a payoff-relevant action maximizing her expected utility given her posterior beliefs. We design polynomial-time querying algorithms that optimize the sender's expected utility in this Bayesian Persuasion game. As a technical contribution, we show that queries form partitions of the space of receiver beliefs that can be used to quantify the sender's knowledge.

Uniformly Strict Equilibrium for Repeated Games with Private Monitoring and Communication

R. McLean, I. Obara, A. Postlewaite

University of Pennsylvania, Rutgers University, UCLA

Cooperation through repetition is an important theme in game theory. In this regard, various celebrated "folk theorems" have been proposed for repeated games in increasingly more complex environments. There has, however, been insufficient attention paid to the robustness of a large set of equilibria that is needed for such folk theorems. Starting with perfect public equilibrium as our starting point, we study uniformly strict equilibria in repeated games with private monitoring and direct communication (cheap talk). We characterize the limit equilibrium payoff set and identify the conditions for the folk theorem to hold with uniformly strict equilibrium.

Consensus and Disagreement: Information Aggregation under (not so) Naive Learning

A. Banerjee, O. Compte

Paris School of Economics

We explore a model of non-Bayesian information aggregation in networks. Agents non-cooperatively choose among Friedkin-Johnsen type aggregation rules to maximize payoffs. The DeGroot rule is chosen in equilibrium if and only if there is noiseless information transmission, leading to consensus. With noisy transmission, while some disagreement is inevitable, the optimal choice of rule amplifies the disagreement: even with little noise, individuals place substantial weight on their own initial opinion in every period, exacerbating the disagreement. We use this framework to think about equilibrium versus socially efficient choice of rules and its connection to polarization of opinions across groups.

Mathematical Finance

Trading with Concave Price Impact and Impact Decay - Theory and Evidence

N. Hey, I. Mastromatteo, J. Muhle-Karbe, K. Webster

École Polytechnique, Imperial College London, Columbia University

We study statistical arbitrage problems accounting for the nonlinear and transient price impact of metaorders observed empirically. We show that simple explicit trading rules can be derived even for general non-parametric alpha and liquidity signals, and also discuss extensions to several impact decay timescales. These results are illustrated using a proprietary dataset of

CFM metaorders, which allows us to calibrate the levels, concavity, and decay parameters of the price impact model and analyze their effects on optimal trading.

Optimal portfolio allocation with uncertain covariance matrix

M. Markov, V. Markov
École Polytechnique

In this paper, we explore the portfolio allocation problem involving an uncertain covariance matrix. We calculate the expected value of the Constant Absolute Risk Aversion (CARA) utility function, marginalized over a distribution of covariance matrices. We show that marginalization introduces a logarithmic dependence on risk, as opposed to the linear dependence assumed in the mean-variance approach. Additionally, it leads to a decrease in the allocation level for higher uncertainties. Our proposed method extends the mean-variance approach by considering the uncertainty associated with future covariance matrices and expected returns, which is important for practical applications.

Risk of Transfer Learning and its Applications in Finance

H. Cao, H. Gu, X. Guo, M. Rosenbaum
École Polytechnique, UC Berkeley

Transfer learning is an emerging and popular paradigm for utilizing existing knowledge from previous learning tasks to improve the performance of new ones. In this paper, we propose a novel concept of transfer risk and analyze its properties to evaluate transferability of transfer learning. We apply transfer learning techniques and this concept of transfer risk to stock return prediction and portfolio optimization problems. Numerical results demonstrate a strong correlation between transfer risk and overall transfer learning performance, where transfer risk provides a computationally efficient way to identify appropriate source tasks in transfer learning, including cross-continent, cross-sector, and cross-frequency transfer for portfolio optimization.

Occupied Processes: Going with the Flow

V. Tissot-Daguette
Princeton University

We develop an Itô calculus for functionals of the "time" spent by a path at arbitrary levels. A Markovian setting is recovered by lifting a process X with its flow of occupation measures $\mathcal{O}$ and call the pair ($\mathcal{O}$,X) the occupied process. While the occupation measure erases the chronology of the path, we show that our framework still includes many relevant problems in stochastic analysis and financial mathematics. The study of occupied processes therefore strikes a middle ground between the path-independent case and Dupire's Functional Itô Calculus. We extend Itô's and Feynman-Kac's formula by introducing the occupation derivative, a projection of the functional linear derivative used extensively in mean field games and McKean-Vlasov optimal control. Importantly, we can recast through Feynman-Kac's theorem a large class of path-dependent PDEs as parabolic problems where the occupation measure plays the role of time. We apply the present tools to the optimal stopping of spot local time and discuss financial examples including exotic options, corridor variance swaps, and path-dependent volatility.

Acknowledgement

The OMI Research Newsletter is curated by Álvaro Cartea, Director of the Oxford-Man Institute, and Fayçal Drissi, postdoctoral researcher at the institute. Paper suggestions are from fellow PhD students at the institute: Álvaro Arroyo, Patrick Chang, Gerardo Durán-Martín, and Marcello Monga.

References

Achitouv, Ixandra, Dragos Gorduza, and Antoine Jacquier. 2023. "Natural Language Processing for Financial Regulation." *arXiv preprint arXiv:2311.08533*.

Altamirano, Matias, François-Xavier Briol, and Jeremias Knoblauch. 2023. "Robust and Conjugate Gaussian Process Regression." *arXiv preprint arXiv:2311.00463*.

Auer, Andreas, Martin Gauch, Daniel Klotz, and Sepp Hochreiter. 2023. "Conformal prediction for time series with Modern Hopfield Networks." *arXiv preprint arXiv:2303.12783*.

Banerjee, Abhijit, and Olivier Compte. 2022. "Consensus and Disagreement: Information Aggregation under (not so) Naive Learning." National Bureau of Economic Research.

Bencomo, Gianluca M, Jake C Snell, and Thomas L Griffiths. 2023. "Implicit Maximum a Posteriori Filtering via Adaptive Optimization." *arXiv preprint arXiv:2311.10580*.

Cao, Haoyang, Haotian Gu, Xin Guo, and Mathieu Rosenbaum. 2023. "Risk of Transfer Learning and its Applications in Finance." *arXiv preprint arXiv:2311.03283*.

Cartea, Álvaro, Patrick Chang, and Gabriel García-Arenas. 2023. "Spoofing and Manipulating Order Books with Learning Algorithms." *Available at SSRN*.

Harris, Keegan, Nicole Immorlica, Brendan Lucier, and Aleksandrs Slivkins. 2023. "Algorithmic Persuasion Through Simulation: Information Design in the Age of Generative AI." *arXiv preprint arXiv:2311.18138*.

Hey, Natascha, Iacopo Mastromatteo, Johannes Muhle-Karbe, and Kevin Webster. 2023. "Trading with Concave Price Impact and Impact Decay-Theory and Evidence." *Available at SSRN*.

Markov, Maxime, and Vladimir Markov. 2023. "Optimal portfolio allocation with uncertain covariance matrix." *arXiv preprint arXiv:2311.07478*.

McLean, Richard, Ichiro Obara, and Andrew Postlewaite. 2023. "Uniformly Strict Equilibrium for Repeated Games with Private Monitoring and Communication." *arXiv preprint arXiv:2311.12242*.

Pinelis, Michael, and David Ruppert. 2023. "Maximizing Portfolio Predictability with Machine Learning." *arXiv preprint arXiv:2311.01985*.

Stok, Robert, and Paul Bilokon. 2023. "From Deep Filtering to Deep Econometrics." *arXiv preprint arXiv:2311.06256*.

Sun, Chuhao, Asaf Cohen, Janes Stokes, and Shravan Veerapaneni. 2023. "Quantum-inspired nonlinear Galerkin ansatz for high-dimensional HJB equations." *arXiv preprint arXiv:2311.12239*.

Tang, Siu Hin, Mathieu Rosenbaum, and Chao Zhou. 2023. "Forecasting Volatility with Machine Learning and Rough Volatility: Example from the Crypto-Winter." *arXiv preprint arXiv:2311.04727*.

Tissot-Daguette, Valentin. 2023. "Occupied Processes: Going with the Flow." *arXiv preprint arXiv:2311.07936*.

Zhang, Boyu, Hongyang Yang, Tianyu Zhou, Muhammad Ali Babar, and Xiao-Yang Liu. 2023. "Enhancing financial sentiment analysis via retrieval augmented large language models." 349–356.